

Book Review: Meinard Müller: Fundamentals of Music Processing

Knox, Don

Published in:
Musicae Scientiae

DOI:
[10.1177/1029864915625414](https://doi.org/10.1177/1029864915625414)

Publication date:
2016

Document Version
Author accepted manuscript

[Link to publication in ResearchOnline](#)

Citation for published version (Harvard):
Knox, D 2016, 'Book Review: Meinard Müller: Fundamentals of Music Processing', *Musicae Scientiae*, vol. 20, no. 2, pp. 269-271. <https://doi.org/10.1177/1029864915625414>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please view our takedown policy at <https://edshare.gcu.ac.uk/id/eprint/5179> for details of how to contact us.

Review of *Fundamentals of Music Processing*, Meinard Müller.

Don Knox

Published in Musicae Scientiae 1029864915625414, first published on January 8, 2016 as doi:10.1177/1029864915625414

In writing *Fundamentals of Music Processing*, Meinard Müller has created an essential handbook for the music information retrieval (MIR) community. There are alternative choices for those seeking to expand their knowledge in this arena, and who are looking for a text which adopts a computational approach to music concepts and sound analysis. For example the recent *Computational Music Analysis* (David Meredith, ed.) by the same publisher. However Müller's book is clearly positioned as a text book, and the author expresses his hope that it will be of use to researchers, lecturers and students of a topic for which he has a clear and infectious enthusiasm. This aim is emphasised by the structure and content of the book, which covers a range of important and popular music processing topics, and strikes a balance between providing succinct and informative background theory and discussing applied examples in the context of real-world problem solving. The author goes as far as proposing several different ways in which the book might be navigated, which he refers to as 'views'. These suggest access points and approaches to navigating the text, and sections which might be of particular relevance to other disciplines. This is invaluable to the target reader – perhaps especially the lecturer who may use this text to plan courses in music analysis and processing. In each chapter, Müller introduces the reader to analysis and processing techniques, algorithms and key underpinning theoretical concepts within the context of applied music analysis and processing tasks.

The first two chapters deal with fundamentals of music representations and sound analysis. A technically oriented view of music is presented which centres around sheet music, symbolic and audio representations. This section encompasses basic music concepts including pitch, chroma and western music notation; MIDI, MusicXML and Optical Music Recognition; time and frequency parameters of audio signals including dynamics, timbre, envelope, modulation and signal spectra. This section concludes with a thorough explanation of the spectrogram, a fundamental concept underpinning music analysis and many of the techniques presented in this book. This succinct and accessible chapter sets the scene for chapter two, which presents a comprehensive treatment of Fourier analysis. This begins with a clear and intuitive overview of Fourier principles, with a view to providing an understanding of frequency analysis crucial to music processing concepts presented in subsequent chapters. This initial overview may be useful to the reader who is looking for a manageable entry point to the topic, but it is quickly followed by a rigorous and thorough treatment of signal and sampling theory, signal space, and the mathematical principles underpinning Fourier analysis in its various computational forms. This chapter in particular is relevant to readers from a number of disciplines, and comprehensively covers the salient topics one would expect in a related Master's or advanced undergraduate degree. By chapter three we have moved on to music processing matters, beginning with the topic of music synchronisation. This chapter presents key concepts of feature extraction and sequence alignment, focussing first on chroma features and their extraction from the spectrogram. The author deals with issues of tuning effects and transposition, useful examples which are typical of the book's focus on practical application of the techniques being discussed. Dynamic time warping is explored in the context of feature matching, applied via

dynamic programming. Potential applications of the technology are explored, including inter-document navigation applications for exploration and synchronised interaction with different interpretations of a musical piece. Chapter four moves on to music structure analysis. In this chapter the context is identification of song form through analysis of the audio representation, and automatic classification of song segments in respect of homogeneity, repetition and novelty. The author deals well with the complexity and nuance inherent in this topic, acknowledging the limitations of the technology and necessary simplification of assumptions required in this context. Key underpinning concepts are explained on the way – including the features characterising song sections; chroma, Mel frequency cepstral coefficients, and measures of rhythm and tempo. The reader is also introduced to self-similarity matrices, crucial to computational music structure analysis as a whole, and also a thorough discussion of enhancement strategies toward improving the usefulness/accuracy of results gained through these methods. This chapter also includes a brief section on audio thumbnailing, the extraction of a segment of audio most representative of a piece of music. Notionally this is most relevant to browsing through large audio/music corpora, but is discussed here in the context of repetition-based structure analysis. Chord recognition is the practical application context for chapter five, which tackles harmonic analysis and briefly touches upon fundamental musical concepts of interval, scales, consonance and dissonance, simple chords and progressions. Müller strikes a great balance between covering the relatively narrow topic of chord labelling, whilst encompassing several key topics salient to other music analysis tasks, for example pattern matching, filtering and feature extraction. There is a particularly good section on hidden Markov models. Again there is some excellent discussion of ambiguities in chord recognition and suggested enhancements to analysis techniques designed to mitigate for this. Tempo and beat tracking is the focus of chapter 6, which presents a solid treatment of this enduring and still-challenging topic for music processing research. Tempo analysis is implemented through creation of tempograms, via Fourier analysis, autocorrelation, and the mid-level cyclic tempogram, which the author posits as the ‘tempo-related counterpart’ to chroma representations. A clear distinction is drawn between tempo analysis and beat tracking, again implemented through dynamic programming. This section encompasses the ‘predominant local pulse’ technique, a novelty-based approach proposed to be sensitive to continuous and sudden tempo variations, and based on the author’s own previous research. Throughout this chapter Müller considers and deals with ambiguity, problems and challenges in tempo and beat tracking, providing useful insight and context to the technological discussion. The final two chapters deal with key challenges for the MIR community, beginning with content based audio retrieval. Chapter 7 focusses on query by example applications, and coheres around the concept of search specificity and granularity. Three specific topics are presented, beginning with audio identification. This essentially means audio fingerprinting, through compact descriptions of sound parameters characterising the audio excerpt in question. This is an application of MIR technology that has gained increasing currency in recent years, through the popularity of applications like Shazam. Discussion covers the key topics of robustness, scalability and the requirement for the most compact representation of data, matching processes, and search strategies based on an indexing and retrieval framework. Next is audio matching, a lower specificity task aimed at identifying pieces of music similar to that identified in the query, but allowing for variations which account for different performances, tempi, and instrumentation. Finally this section turns to version identification, a global similarity task comparing whole tracks. This section features some excellent discussion on what different ‘versions’ might actually mean – medleys, mixes, remixes, remasters, and mashups for example, and the nature of these modifications in comparison

to the 'essence' of the track we are searching for. The final chapter of the book deals with the 'grand challenge' of music source separation. Three scenarios are considered: harmonic/percussive separation, melody separation, and nonnegative matrix factorisation (NMF) to produce a note-wise decomposition of the music signal. Discussion centres around source separation, and the recognition that achieving this goal to a consistently high quality is still an unsolved problem. Along the way several key topics are discussed in considerable detail, including spectrogram filtering and masking, signal reconstruction from the modified short-time Fourier transform, predominant F0 estimation following, harmonic summation, salience representations, and discussion on key applications such as query by humming and score-informed audio editing.

The stated aim of this book is to introduce readers to MIR, exploring well-established applications as a context to explain salient algorithms, concepts and techniques in depth. In this regard it is a great success. Considerable effort and thought has gone into the structure and content of this book, to the benefit of those who will use it. Each chapter ends with a summary of key points discussed in the context of the wider literature, practical implementations and limitations, followed by a list of exercises toward supporting the reader's understanding of each topic area. It's worth noting that the solutions to the exercises are not contained in the book, rather they are provided via an accompanying website. Another of the author's aims is to make the book relevant and accessible to students from varied backgrounds and study areas. Presenting technological concepts in the context of practical applications, and the liberal use of excellent, clear and informative diagrams has undoubtedly resulted in a broadly accessible book. However, as one would expect in a text adopting a computational perspective, there is rigorous mathematical treatment of all the key processes and concepts covered in the book. For readers from other disciplines, who perhaps have less of a firm background in signal processing terminology and notation, this may present somewhat of an obstacle to engaging with some parts of the text. But overall Müller strikes an excellent balance between the mathematics, intuitive descriptions and illustrations. The mathematical treatment of processes and algorithms will be invaluable to many students, researchers and tutors exploring the world of computational music analysis. The scope of the book is well pitched in terms of the research topics it covers, although perhaps the range of music features discussed could have been broadened to encompass the large variety of music and sound parameters employed in MIR endeavours. Overall though, this is a very impressive book, and an excellent course in music processing. This is essential reading for the MIR community.

Reference

Meredith, D. (2015). *Computational Music Analysis*. Springer.